MIT Open Access Articles

Automated compact dynamical modeling:
An enabling tool for analog designers

Massachusetts Institute of Technology

Automated Compact Dynamical Modeling: An Enabling Tool for Analog Designers

Bradley N. Bond
Massachusetts Institute of Technology
bnbond@mit.edu

Luca Daniel
Massachusetts Institute of Technology
luca@mit.edu

ABSTRACT

In this paper we summarize recent developments in compact dynamical modeling for both linear and nonlinear systems arising in analog applications. These techniques include methods based on the projection framework, rational fitting of frequency response samples, and nonlinear system identification from time domain data. By combining traditional projection and fitting methods with recently developed convex optimization techniques, it is possible to obtain guaranteed stable and passive parameterized models that are usable in time domain simulators and may serve as a valuable tool for analog designers in both top-down and bottom-up design flows.

Categories and Subject Descriptors: J.6 [Computer Aided Engineering]: Computer-aided design (CAD), G.1.6 [Optimization]: Convex programming, I.6.0 [Simulation and Modeling]: General
General Terms: Algorithms
Keywords: Model reduction, Analog design, Semidefinite programming, Parameterized modeling, Compact modeling

1. INTRODUCTION

As feature sizes keep decreasing and operating frequencies keep increasing, the performance of many complex analog components and systems is becoming crucially dependent on what the designers identify as "non-idealities" or "second order effects". These are phenomena such as non-linearities, process variations, skin effects, proximity effects, and substrate and magnetic coupling. Analog designers would therefore greatly benefit from tools that generate automatically compact dynamical models which capture *quantitatively* all such second order effects. Digital designers have indeed been already benefitting for decades from model order reduction tools to speed-up circuit simulations within the context of a bottom-up library characterization step for timing analysis. However, several significant differences between analog and digital applications limit severely the usability of the existing model order reduction techniques.

The first important difference is that, within the digital applications, the systems to be "reduced" are mainly, if not exclusively, RLC interconnect networks generated by parasitic extractors (e.g. resistance, capacitance and inductance extractors). Such systems are linear, and are always described by *constant, structured, and*

semidefinite matrices. For linear systems with such specific properties, a body of methods exist (e.g. [1, 2]) that can generate automatically compact stable and passive dynamical models suitable for time domain circuit/gate level simulations. On the contrary, when working with analog applications, one needs to handle *nonlinear* circuit blocks or microelectromechanical (MEMS) devices. Even when considering only the linear part of the analog systems (e.g. RF inductors, passive power combiners, transmission lines) the constraint of "constant semidefinite matrices" is always inevitably violated. This is because the electromagnetic field solvers used to analyze analog structures inevitably introduce numerical discretization errors at high frequencies, and because they are forced by the large size of the problems to use advanced formulations and approximation techniques such as substrate Green functions and fast matrix-vector products. In many cases no internal description of the systems is even available, and the only information provided by the field solvers, or by measured data, is a set of frequency response values (if the system is linear), or a set of time domain input-output signals (if the system is nonlinear).

Another important difference between digital and analog systems is that in digital applications higher level system simulators (e.g. timing analyzers) or top-down design methodologies do *not* require dynamical state-space models for each gate, but rather just a small set of simple equations or tabulated values relating performance (or behavioral) parameters. However, in order to enable system level simulation and fast tradeoff explorations for analog applications, leading to robust designs insensitive to process variations and to all high frequency second order effects, it is essential to have the ability to instantiate instantaneously parameterized compact *dynamical state-space models* of circuit blocks.

To better highlight the differences between digital and analog applications, we would like to emphasize that the classical name, "model order reduction", used in the digital community, is completely inadequate to describe the variety of techniques that are being developed to handle analog systems. We propose instead the name "automated compact dynamical modeling". Here the term *dynamical* indicates state space models, emphasizing the fact that for system level analog simulations it is not sufficient to generate simple static behavioral models relating performance parameters of a circuit block as in digital applications. The term *compact* as opposed to "reduced order", emphasizes the fact that model complexity is not strictly tied to the model order (i.e. the number of equations) as is the case in the digital interconnect linear system world. The complexity of the generated system should instead be measured in terms of how efficiently the generated model can be employed in a time domain simulator. When working with analog nonlinear systems one quickly realizes that small model orders do not necessarily imply fast simulations. Finally, although in some cases one might want to "reduce" a large model already given in dynamical state space form, in all the other cases typical of the analog world, one is only given frequency domain data, or input-

output time domain data. Therefore a compact dynamical model might need to be "'identified" or "fitted" as opposed to "reduced". As a matter of fact, in order to address the many new challenging analog constraints, during the last few years the automated compact dynamical modeling community has been forced to move beyond traditional projection-based model order reduction approaches, embracing for instance recent developments in convex optimization. In some cases, some of the more traditional projection approaches or unconstrained fitting techniques have been adapted to the analog world by simply adding a perturbation step that enforces properties such as stability and passivity as a post-processing operation. In other cases, completely new techniques have been, and are still currently being invented.

The remainder of this paper is organized as follows. Section 2 presents projection methods for stable and passive modeling of linear and nonlinear systems described in state-space form. Section 3 surveys methods for compact modeling of linear systems using rational approximation techniques for transfer matrix fitting, combined with optimization-based stability and passivity enforcement. Section 4 summarizes recent developments in stable dynamical model fitting techniques using optimization to enforce stability in nonlinear dynamical systems. Lastly, Section 5 discusses what further steps need to be taken, both by model reduction researchers and by the rest of the EDA community, to help compact dynamical modeling become a mainstream tool supporting analog design.

2. PROJECTION-BASED METHODS

2.1 The Traditional Projection Framework

Most of the traditional model order reduction techniques can be interpreted within a projection framework. In such a framework the solution to a given large linear multi-input multi-output system

$$E\dot{x} = Ax + Bu, \qquad y = C^T x, \qquad (1)$$

is approximated in a low-dimensional space $x \approx V\hat{x}$, where $V \in \mathbb{R}^{N \times q}$ is the right projection matrix, $\hat{x}$ is the reduced state vector, and $N >> q$. A reduced set of equations is then obtained by forcing the residual, $r(V\hat{x}) = EV\dot{\hat{x}} - AV\hat{x} - Bu$, to be orthogonal to the subspace defined by a left projection matrix U, i.e. $U^T r(V\hat{x}) = 0$. The resulting state-space model has the form

$$\hat{E}\dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}u, \qquad y = \hat{C}^T\hat{x} \qquad (2)$$

where $\hat{E} = U^T E V$, $\hat{A} = U^T A V$, $\hat{B} = U^T B$, and $\hat{C} = V^T C$. The accuracy of the reduced model created via projection is completely determined by the choice of projection matrices U and V. The most common approaches for selecting the vectors are methods based on balanced truncation, moment matching, and singular value decomposition (e.g proper orthogonal decomposition, principal components analysis, or Karhunen-Loeve expansion). For more details on generating projection vectors see [2, 3].

When applying projection to a nonlinear system of the form

$$\dot{q}(x) = f(x, u), \qquad y = g(x), \qquad (3)$$

the resulting reduced model can be expressed as

$$\dot{\hat{q}}(\hat{x}) = \hat{f}(\hat{x}, u), \qquad y = \hat{g}(\hat{x}) \qquad (4)$$

where $\hat{q}(\hat{x}) = U^T q(V\hat{x})$ and $\hat{f}(\hat{x}, u) = U^T f(V\hat{x}, u)$. In this case the reduced order model (4) does not lead to substantially faster simulation times when used in a time domain integrator, because the vector field term $f(V\hat{x}, u)$ in the reduced model still has $O(N)$ complexity. To solve this critical problem it is common to apply projection not to the original nonlinear system (3), but instead to a nonlinear system defined by carefully selected approximate nonlinear functions $\tilde{q}(x)$ and $\tilde{f}(x, u)$.

For weakly nonlinear systems it is common to use a polynomial expansion for the approximation. For example, for a simplified nonlinear system $\dot{x} = f(x) + Bu$, the vector field $f(x)$ is approximated as

$$f(x) \approx \tilde{f}(x) = A_0 + A_1 x + A_2(x \otimes x) + \dots . \qquad (5)$$

where A_k is an $N \times N^k$ matrix containing the kth derivatives of $f(x)$. The new vector field $\tilde{f}(x)$ can be easily projected given the Kronecker product property $(V\hat{x}) \otimes (V\hat{x}) = (V \otimes V)(\hat{x} \otimes \hat{x})$, resulting in the reduced vector field

$$\hat{f}(\hat{x}) = U^T \tilde{f}(V\hat{x}) = \hat{A}_0 + \hat{A}_1 \hat{x} + \hat{A}_2(\hat{x} \otimes \hat{x}) + \dots$$

where $\hat{A}_0 = U^T A_0$, $\hat{A}_1 = U^T A_1 V$ and $\hat{A}_2 = U^T A(V \otimes V)$. Analysis of polynomial reduced models can be further simplified using ideas from Volterra theory, as was done in [4, 5].

An alternative approach, better suited for highly nonlinear systems, consists of interpolating between a collection of local models

$$f(x) \approx \tilde{f}(x) = \sum_i w_i(x)\tilde{f}_i(x), \qquad (6)$$

where $w(x)$ are weighting functions. The local models are chosen so that each individual model $\tilde{f}_i(x)$ can be reduced using projection

$$\hat{f}(\hat{x}) = U^T \tilde{f}(V\hat{x}) = \sum_i w_i(\hat{x})U^T \tilde{f}_i(V\hat{x}), \qquad (7)$$

such that $U^T \tilde{f}_i(V\hat{x})$ has low complexity. For instance, $\tilde{f}_i$ could be polynomial functions, or even linear $\tilde{f}_i(x) = A_i x + k_i$, in which case $U^T \tilde{f}_i(V\hat{x}) = U^T A_i V\hat{x} + U^T k_i$. The key idea behind the Trajectory PieceWise Linear (TPWL) method [6] and related techniques [7, 8] is to generate the local models along solution trajectories of (3) in response to typical training inputs, thus keeping the required number of local models small. A more detailed survey of these methods can be found in [9].

2.2 Stable Projection for Linear Systems

Traditionally it is assumed that the original large system (1) possesses an extremely special structure: $E = E^T \succeq 0$, $A \preceq 0$, and $B = C$. In such cases selecting $U = V$ (known as a congruence transform or Galerkin projection) will preserve stability and passivity in the reduced system for *any choice of* V. While all digital RLC type interconnect networks possess the required semidefinite structure, for analog modeling it is unfortunately completely unrealistic to restrict consideration to only semidefinite systems. Therefore for the vast majority of analog systems, the congruence transform cannot guarantee stability and passivity of the generated model. One possible computationally cheap solution is to use as a first step any of the available traditional projection based methods (including congruence transforms) and then attempt to perturb the generated model to enforce stability and passivity. One semidefinite formulation of this problem is

$$\min_{\Delta\hat{E}, \Delta\hat{A}, \Delta\hat{C}} ||\Delta\hat{E}|| + ||\Delta\hat{A}|| + ||\Delta\hat{C}|| \quad \text{subject to} \qquad (8)$$

$$\hat{E} \succeq 0, \quad \hat{A} + \hat{A}^T \preceq 0, \quad \hat{B} = \hat{C}.$$

where $\hat{E} = U^T E V + \Delta\hat{E}$, $\hat{A} = U^T A V + \Delta\hat{A}$, $\hat{B} = U^T B$, and $\hat{C} = V^T C + \Delta\hat{C}$. Here stability and passivity are enforced in the reduced model by forcing it to be described by semidefinite system matrices, which introduces no loss of generality even if the original system (1) is not described by semidefinite matrices [10].

Unfortunately, in most cases any such perturbation could completely destroy the accuracy of the reduced model. Instead of perturbing the reduced model, a better approach that can guarantee accuracy in the reduced model is to perturb one of the projection matrices. That is, given U and V, search for a "small" ΔU such

that system (2), defined by reduced matrices $\hat{E} = (U + \Delta U)^T EV$, $\hat{A} = (U + \Delta U)^T AV$, $\hat{B} = (U + \Delta U)^T B$, and $\hat{C} = V^T C$ is passive. This problem can similarly be formulated as a semidefinite program

$$\min_{\Delta U} ||\Delta U|| \quad \text{subject to} \tag{9}$$

$$\hat{E} \succeq 0, \quad \hat{A} + \hat{A}^T \preceq 0, \quad \hat{B} = \hat{C}.$$

It can be shown that if the original model (1) is stable and passive, then for any projection matrix V there exist projection matrices U such that the resulting reduced model is stable and passive [10]. These approaches can also be used to stabilize unstable models of stable physical systems, where the instability can arise either from numerical errors in the field solver analysis of the system, or from a first stage projection.

2.3 Stable Vector-Field Approximation

Guaranteeing stability and passivity in the generated compact model becomes more difficult when combining projection with function approximation for nonlinear systems. In general function approximations do not preserve stability, i.e. the nonlinear system described by $\tilde{q}, \tilde{f}$ (e.g. (5) or (6)) is in general *unstable*. For instance, a simple quadratic model, such as $\dot{x} = -x^2 - x$, can never have a globally stable equilibrium point. Although global stability may be unattainable in such cases, it is almost always possible to enforce local stability in the reduced model. Suppose $x = 0$ is a stable equilibrium for the nonlinear system described by q, f. If $\tilde{q}, \tilde{f}$ are chosen such that both systems have the same linear approximation around $x = 0$, as is the case when using polynomial and piecewise polynomial approximations, then $x = 0$ is guaranteed to be locally stable for the approximate system. With this information, it is possible to use the stability-preserving projection techniques discussed in Section 2.2 to select U, V such that a linearization of the reduced nonlinear system $\hat{q}, \hat{f}$ has a locally stable equilibrium point $x = 0$. Furthermore, it is possible to enforce local stability at multiple linearizations by employing a nonlinear left-projection matrix [11].

2.4 Parameterization of Projection Methods

Generating a parameterized reduced model, such as $\hat{E}(p)\dot{\hat{x}} = \hat{A}(p)\hat{x} + \hat{B}(p)u$ for a linear system where p is a vector of design parameters, is of critical importance if the models are to be used for system level design trade-off explorations or for robustness verification in the presence of process variations. Two modifications to the previously described projection procedures must be made when constructing parameterized models. First, the subspace defined by V must capture the solution response to changes in parameter values. Expanding the subspace is typically achieved for linear or linearized systems by generating projection vectors that match the frequency response derivatives with respect to the parameters p in addition to the frequency [12, 13, 14]. Alternative approaches for handling variability resulting from a large number of parameters are based on sampling and statistical analysis [15, 16].

The second issue involves specifically the case of nonlinear parameter dependence, where the system matrix or vector field must be able to cheaply capture changes in p. One way to make a parameterized system matrix $A(p)$, or vector field, projectable with respect to the parameters is to represent them as a sum of non-parameterized functions that are linear in scalar functions of the original parameters. For instance, for the parameterized linear system $\dot{x} = A(p)x$ we seek to approximate and project as follows

$$A(p) \approx \sum_{i=0}^{\kappa} \tilde{A}_i g_i(p) \longrightarrow \hat{A}(p) = \sum_{i=0}^{\kappa} (U^T A_i V) g_i(p) \tag{10}$$

such that $\hat{A}_i = U^T A_i V$ are constant matrices and can be precom-

puted. Here $g_i(p)$ are scalar functions of the original parameter set p. The matrix approximation in (10) can be achieved using a polynomial expansion if $A(p)$ is known analytically, or via fitting in the case when only samples of the matrix $A_k = A(p_k)$ are known [14]. Similar approaches can be used to generate parameterized nonlinear models [17].

3. RATIONAL TRANSFER MATRIX FITTING

Projection methods have been extremely successful for certain classes of linear systems, but in many applications, such as when modeling analog passive components affected by full-wave effects or substrate effects, the resulting system matrices include delays or frequency dependency. Guaranteeing stability and passivity in such models through projection is an extremely challenging task. An alternative class of methods capable of generating stable and passive models for such linear systems are based on transfer matrix fitting, where the transfer matrix, or frequency response, is defined as $H(s) = C^T(sE - A)^{-1}B$.

A multi-port linear system representing impedance or admittance is stable when its transfer matrix has no poles with positive real part and any pole on the imaginary axis is simple. The system is passive if, in addition to stability it satisfies the positivity condition:

$$H(j\omega) + H(jw)^\dagger \succeq 0 \quad \forall \omega \tag{11}$$

Sometimes stability and positivity are collectively referred to as positive realness.

Schemes for enforcing passivity (i.e. stability and positivity) in transfer matrix fitting methods can be broadly classified into two approaches: those which use unconstrained minimization combined with post-processing perturbation to enforce stability and positivity; and those that simultaneously enforce stability and positivity during the fitting process.

3.1 Unconstrained Rational Fitting

Given a set of frequency response samples $\{H_i, \omega_i\}$, where $H_i = H(j\omega_i)$ is the transfer matrix of some unknown multiport linear system, the compact modeling task is to construct a low-order rational transfer matrix $\hat{H}(s)$ such that $\hat{H}(j\omega_i) \approx H_i$. The most common formulation of this problem is an L_2 minimization of the sum of squared errors

$$\min_{\hat{H}} \sum_i ||H_i - \hat{H}(j\omega_i)||^2. \tag{12}$$

Even ignoring for a moment stability and passivity constraints, the unconstrained minimization problem (12) is still non-convex (because $\hat{H}$ is a rational function) and therefore extremely difficult to solve. Direct minimization using a nonlinear least-squares algorithm, such as Levenberg-Marquardt as in [18], yields only a locally optimal result.

Rather than solving non-convex minimization problems, many alternative methods apply a relaxation to the objective function resulting in an optimization problem that can be solved efficiently. One very popular approach, referred to as *vector fitting*, relaxes the problem by first finding the poles a_k with an iterative procedure, and then, with the poles fixed, solves a linear least-squares problem to determine the residue matrices R_k [19], resulting in a transfer function in pole-residue form

$$H(s) = \sum_k \frac{R_k}{s - a_k} + D. \tag{13}$$

An alternative relaxation considers the transfer function in numerator and denominator form, $H(s) = P(s)/q(s)$, where $P(s)$ is a matrix of polynomials and $q(s)$ is a scalar polynomial. By first identifying the real parts of the numerator $B(j\omega) = \Re P(j\omega)$ and denominator $a(j\omega) = \Re q(j\omega)$, the problem can be transformed

into a second order cone programming problem, which is a special case of semidefinite programming and can thus be solved in an efficient manner [20, 21]

$$\min_{B,a} \sum_i \frac{|a(j\omega_i)\Re H_i - B(j\omega_i)|^2}{a(j\omega_i)}. \tag{14}$$

For stable systems, P and q can then be uniquely determined from B and a via inverse Hilbert transform after fitting.

3.2 Stability and Passivity via Post-Processing

In vector fitting methods, stability is achieved by simply reflecting across the imaginary axis all the unstable poles obtained during the iterative fitting procedure [22]. Once a stable model is obtained, passivity is obtained by further enforcing positivity condition (11), which is a positive-definiteness constraint on H and therefore implies that the transfer matrix $H(j\omega)$ has no eigenvalues with negative real part at any frequency ω. Efficient algebraic methods based on computing the eigenvalues of the Hamiltonian matrix for the state-space model can be used to determine the set of positivity violations [23]

$$\Omega = \left\{ \omega_k \Big| \Re\{\lambda(\hat{H}(j\omega_k))\} < 0 \right\}.$$

During the post-processing step, pole locations are kept fixed and passivity is obtained by altering only the residues. This is achieved by transforming from pole-residue form (13) to state-space form (1) and perturbing only the output vector C. A generic formulation of the positivity-enforcing minimal perturbation can be stated as

$$\min_{\Delta C} ||\Delta C||_X, \quad \text{subject to } \mathcal{P}(\Omega) \tag{15}$$

where $\mathcal{P}(\Omega)$ is a positivity constraint for the transfer matrix over the set of positivity violations, and $||\cdot||_X$ denotes the norm used for quantifying the effects of the perturbation on the accuracy of the model. In [24] it was proposed to select the norm that produces the minimal perturbation to the impulse response of system (1), defined as $||\Delta C||_X = ||\Delta C K^T||_2$ where $K^T K = W$ is the controllability Grammian satisfying $EWA^T + AWE^T + BB^T = 0$. Other possible choices for the objective function are presented in [23]. The positivity constraint $\mathcal{P}(\Omega)$ is enforced using first-order perturbation to the eigenvalues of the Hamiltonian matrix, and can be expressed as a linear matrix equality or matrix inequality in terms of the unknown perturbation vector ΔC.

3.3 Stability and Passivity During Fitting

In an alternative kind of approach (typically corresponding to the use of relaxation (14)), stability and passivity are enforced as constraints *during* the fitting procedure. Specifically, stability is achieved by requiring that $a(j\omega) > 0$ is a strictly positive polynomial, where $a(j\omega)$ is the real part of the denominator $q(j\omega)$ of the transfer matrix $\hat{H}(j\omega)$. This constraint is efficiently enforced using a sum-of-squares (SOS) relaxation, which provides a convex relation to the non-convex problem of verifying global positivity of a multivariate polynomial. To additionally obtain passivity, the positivity constraint (11) requires the real part of the numerator $B(j\omega) = B(j\omega)^T \succeq 0$, to be a symmetric positive semidefinite matrix function. Such constraint can also be enforced using a SOS relaxation, resulting in a single efficient semidefinite optimization problem [20, 21].

3.4 Parameterized Rational Fitting

There are two possible approaches to generating a parameterized transfer matrix $\hat{H}(s, p)$ from a given set of frequency response samples and parameter values $\{H_i, \omega_i, p_i\}$. The first approach is to fit simultaneously to the frequency response data and parameter values, i.e. minimizing the sum of $|\hat{H}(j\omega_i, p_i) - H_i|^2$. This approach was first proposed in [20] along with simultaneous enforcement of stability and passivity. Alternatively, one may identify a non-passive model $\hat{H}(s, p)$ and utilize the perturbation ideas from section 3.2 to obtain stability and passivity at any desired parameter value p.

The alternative approach is to first fit a set of non-parameterized transfer matrices $\{\hat{H}_k(j\omega)\}$ such that $\hat{H}_k(j\omega_i) \approx H_i$ corresponding to the parameter set p_k, and then "interpolate" between the set of identified models to obtain a parameterized transfer matrix satisfying $\hat{H}(j\omega, p_k) \approx \hat{H}_k(j\omega)$. Techniques using this approach have enforced stability and passivity via post-processing perturbation as described in section 3.2 [25], but no such techniques at this time are capable of enforcing stability and passivity as constraints during interpolation, as described in section 3.3.

4. STABLE DYNAMICAL MODEL FITTING

The success of projection-based model reduction techniques for linear systems led to the natural extension of such techniques to nonlinear systems. However, nonlinear systems arising in analog applications are often relatively small $(10 - 1000$ equations) compared to their linear counterparts $(10^6$ equations), and the explicit nonlinear state-space equations (3) are frequently unavailable in a simple manner. An alternative to projection approaches for nonlinear systems is what we refer to as *dynamical model fitting*.

4.1 Dynamical Model Fitting

Dynamical model fitting is the task of fitting a *dynamical* state-space model (as opposed to fitting a "static" function such as a transfer matrix) to a given set of data pairs $\{\tilde{t}_i, \tilde{u}_i, \tilde{y}_i\}$ obtained from the original, possibly unknown, system. In this section for simplicity we consider only input-output models of the form

$$\dot{y} = \hat{f}(y, u), \tag{16}$$

but dynamical model fitting techniques are capable of identifying more general state-space models of the form (4) (see [26] for a thorough description of the state-space case). Ideally the identified model should minimize the sum of the *output errors* $|y_i - \tilde{y}_i|^2$, where $y_i = y(\tilde{t}_i)$ is the output of the identified model (16) in response to given input samples $u(\tilde{t}_i) = \tilde{u}_i$. Minimization of the exact output error is a difficult problem because in general there is no explicit expression for the model output y in terms of the unknown model coefficients. A simpler problem is to search for $\hat{f}$ in order to minimize instead the *equation error*

$$\min_{\hat{f}} \sum_i ||\dot{\tilde{y}}_i - \hat{f}(\tilde{y}_i, \tilde{u}_i)||_2^2. \tag{17}$$

If the unknown functions are defined using a set of predetermined basis functions

$$\hat{f}(y, u) = \sum_j \alpha_j \phi(y, u), \tag{18}$$

where ϕ could be for instance polynomial functions of y and u, then an unconstrained minimization problem can be solved for the unknown coefficients α using standard least-squares [27].

4.2 Stable Dynamical Fitting

As seen in the case of transfer matrix fitting, unconstrained minimization generally produces unstable models, and therefore stability must be enforced as a constraint during model identification. Unfortunately, enforcing stability for nonlinear dynamical systems is a difficult task. Traditional identification methods either completely ignore stability constraints or impose very restrictive structures on the model vector field $\hat{f}$ [27]. In the most general sense, stability or passivity of a dynamical system can be certified through the existence of a positive definite storage function $L(y) \succ 0$ that

satisfies a dissipation constraint $\dot{L}(y) \leq \sigma(y, u)$ for some supply rate $\sigma(y, u)$ (see [28] for more details). For example, if we consider the supply rate $\sigma(y, u) = |u|^2 - |y|^2$ (corresponding to bounded input bounded output stability) and the storage function $L(y) = y^T y$, then system (16) is globally stable if the dissipation constraint

$$y^T \hat{f}(y, u) + \hat{f}(y, u)^T y + y^T y - u^T u \leq 0 \qquad (19)$$

is satisfied for all possible y and u satisfying (16). Therefore, one approach to identifying stable nonlinear models is to solve optimization problem (17) subject to stability constraint (19). It is important to note that dissipation constraint (19) is just one possible constraint that certifies stability of the nonlinear system (other constraints can be obtained by considering different storage functions) and is thus a sufficient condition for stability *but not a necessary condition*. The precise formulation of the stability constraints is an important consideration because, even if $\hat{f}$ is considered to be fixed, verifying whether or not constraint (19) is satisfied is a difficult task. As a result, it is typically necessary to restrict the class of functions $\hat{f}$ considered to those for which inequalities such as (19) can easily be globally verified. For example, if $\hat{f}(y, u)$ is chosen as a polynomial function, then constraint (19) can be enforced as a semidefinite constraint using the sum-of-squares relaxation previously described in section 3.3.

In many cases enforcing dissipation constraints, such as (19), globally for all possible solutions y, u is restrictive, since many stable systems do not satisfy such sufficient (but not necessary) constraint. An alternative suboptimal approach is to consider instead only local stability, which corresponds to linearizations of the nonlinear system (16) being stable. Enforcing local stability at a finite number of points is a much simpler task because it transforms a single nonlinear stability constraint such as (19) into a collection of simple linear constraints, such as searching for positive-definite matrices $P_i = P_i^T$ satisfying $P_i A_i + A_i^T P_i \leq 0$, where $A_i = \partial \hat{f}(y, u)/\partial y$ evaluated at some y_i and u_i, which is a semidefinite constraint. One reasonable choice for solutions around which to enforce local stability is the given set of training data samples $\tilde{u}_i, \tilde{y}_i$ [26].

4.3 Bounding Output Error

Minimization of equation error in optimization problem (17), subject to stability constraint, is a relatively ad-hoc procedure. As a matter of fact, a small equation error does not guarantee a small output error, which is the true desired measure of accuracy for the model. This problem is analogous to the linear algebra notion that for some $b_1 = Ax_1$ and $b_2 = Ax_2$, it is possible that $||b_1 - b_2||$ is small when $||x_1 - x_2||$ is large.

By imposing certain constraints on the model to be identified it is possible to relate the equation error to the output error for nonlinear systems. For instance, it was shown in [29] that if the identified model is *incrementally* stable, then the equation error serves as an upper bound for the output error. Incremental stability, which requires that perturbations to solutions decay, is a strong notion of stability that implies traditional weaker notions such as bounded input bounded output stability. Although enforcing incremental stability as proposed in [29] is restrictive, recent developments in [26] provide a relaxed constraint allowing for identification of a larger class of stable systems with output error bound at a computationally cheaper cost.

4.4 Parameterized Model Fitting

As was the case for parameterized transfer matrix fitting in Section 3.4, parameterized nonlinear models $\dot{y} = \hat{f}(y, u, p)$ can be constructed by either fitting simultaneously to input-state-output data and parameter data, or by first fitting a collection of non-parameterized models, and then interpolating between them. For the first approach, the basis functions must be dependent on the parameters as well as the input and output, e.g. $\psi_i = \psi_i(y, u, p)$ in (18). In the simplest case the parameters can be treated as extra constant inputs, in which case no alterations to the optimization procedure need to be made [26]. Alternatively, one can identify separate nonlinear models $\dot{y} = \hat{f}_k(y, u)$ for each set of parameter values and attempt to identify a single model that "interpolates" the individual models such that $\hat{f}(y, u, p_k) \approx \hat{f}_k(y, u)$ and such that stability is enforced. To our knowledge there is currently no reliable method available for achieving this goal.

5. CONCLUSIONS & FUTURE DIRECTIONS

In terms of the specific techniques at the core of compact dynamical modeling, we believe that the future is neither in projection methods nor in fitting techniques, but rather in a combination of the two. On one side, projection methods are restrictive because they require explicit knowledge of system equations, and in general can preserve stability and passivity only when such equations have a very specific structure. On the other side, fitting techniques can be used to force stability and passivity without knowledge of the original system equations, but they are at a fundamental disadvantage because they are working with a limited amount of information about the original system. We believe therefore that if information about system equations and projection vectors is available, it should be incorporated also into fitting techniques.

Leaving aside for a moment the discussion on the specific core techniques, we also believe that the following general principles should guide the research in automated compact dynamical modeling, in order to finally turn it into a reliable and effective tool in the hands of analog designers.

The most important principle is that the generated dynamical models should *always* be certifiably reliable (i.e. stable and passive) for time domain simulation. Since stability and passivity are *global* properties of the systems, research should be strongly driven toward algorithms that can certify such properties for *any* kind of possible input signals, and not just for a subset of signals of a special form, or within a limited frequency band of interest to the user.

Although models of complex analog systems are unlikely to ever be globally accurate for any kind of input signals, fortunately, global accuracy is *not* as necessary as global stability and passivity. Nevertheless, we do believe that research should be driven toward algorithms that clearly identify the specific subsets of input signals for which the generated models are certifiable accurate, and for which error bounds can be provided.

In order to truly enable high level system verification and effective design tradeoff exploration, research efforts should foster algorithms that generate compact dynamical models which can then be instantiated instantaneously by the analog design users for a range of physical and design parameters. As for the case of input signals, we believe that global accuracy is *not* necessary for the parameters. Nevertheless, research should be driven toward algorithms that clearly identify the specific continuous range of parameter values for which the models can be instantiated with certifiable accuracy in the form of error bounds.

From a practical point of view, the generated models must also be easily usable within existing commercial circuit simulators and high level simulators. On one side, algorithm developers should focus on generating dynamical models that can be described by existing interfaces (e.g. by explicitly and automatically generating Verilog-A modules that implement dynamical state space models). On the other side, EDA circuit simulators should try and provide even more direct and efficient interfaces.

Finally, we recognize that implementing or developing new modeling techniques can require a large initial investment of time and effort, which may deter its wide spread to many potential interesting applications not only within the analog community, but also to

other communities including modeling of devices, nanotechnologies, biological systems and biomedical systems. In order to facilitate dissemination and future collaboration across several compact dynamical modeling communities from different fields, algorithm developers, whenever possible, should try and provide public domain distributions of software and scripts implementing their approaches, including examples and test cases.

APPENDIX A: Optimization Software.

All of the optimization problems formulated in this paper can be solved using the freely available software SPOT [30] and SeDuMi [31], and Matlab tools implementing methods described in this paper can be found online at [32, 33, 34].

6. REFERENCES

[1] A. Odabasioglu, M. Celik, and L. T. Pileggi. PRIMA: passive reduced-order interconnect macromodeling algorithm. *IEEE Trans. Computer-Aided Design*, 17(8):645–654, August 1998.

[2] A. C. Antoulas, D. C. Sorensen, and S. Gugercin. A survey of model reduction methods for large-scale systems. *Structured Matrices in Operator Theory, Numerical Analysis, Control, Signal and Image Processing*, 2001.

[3] S. Gugercin and A. Antoulas. A survey of model reduction by balanced truncation and some new results. *International Journal of Control*, 77(8):748–766, May 2004.

[4] J.R. Phillips. Projection-based approaches for model reduction of weakly nonlinear, time-varying systems. *IEEE Trans. Computer-Aided Design*, 22(2):171–87, 2003.

[5] P. Li and L. T. Pileggi. Norm: Compact model order reduction of weakly nonlinear systems. In *Proc. of the ACM/IEEE Design Automation Conference*, June 2003.

[6] M. Rewienski and J. White. A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices. *IEEE Trans. Computer-Aided Design*, 22(2):155–70, Feb 2003.

[7] S. Tiwary and R. Rutenbar. Faster, parametric trajectory-based macromodels via localized linear reductions. In *Proc. of IEEE/ACM International Conference on Computer Aided-Design*, November 2006.

[8] N. Dong and J. Roychowdhury. General purpose nonlinear model-order reduction using piecewise-polynomial representations. *IEEE Trans. Computer-Aided Design*, 27(2):249–264, Feb 2008.

[9] R.A. Rutenbar, G.G.E. Gielen, and J. Roychowdhury. Hierarchical modeling, optimization, and synthesis for system-level analog and rf designs. *Proceedings of the IEEE*, 95(3):640–669, march 2007.

[10] B. Bond and L. Daniel. Guaranteed stable projection-based model reduction for indefinite and unstable linear systems. In *Proc. of IEEE/ACM International Conference on Computer Aided-Design*, San Jose, CA, 2008.

[11] B. Bond and L. Daniel. Stable reduced models of nonlinear descriptor systems through piecewise linear approximation and projection. *IEEE Trans. Computer-Aided Design*, 28(10):1467–1480, Oct 2009.

[12] D. S. Weile, E. Michielssen, Eric Grimme, and K. Gallivan. A method for generating rational interpolant reduced order models of two-parameter linear systems. *Applied Mathematics Letters*, 12:93–102, 1999.

[13] P. Gunupudi, R. Khazaka, and M. Nakhla. Analysis of transmission line circuits using multidimensional model reduction techniques. *Advanced Packaging, IEEE Transactions on*, 25(2):174–180, may 2002.

[14] L. Daniel, C. S. Ong, S. C. Low, K. H. Lee, and J. K. White. A multiparameter moment matching model reduction approach for generating geometrically parameterized interconnect performance models. *IEEE Trans. Computer-Aided Design*, 23(5):678–93, May 2004.

[15] Z. Zhu and J. Phillips. Random sampling of moment graph: a stochastic krylov-reduction algorithm. In *DATE '07: Proc. of the conference on Design, automation and test in Europe*, pages 1502–1507, 2007.

[16] T. El-Moselhy and L. Daniel. Variation-aware interconnect extraction using statistical moment preserving model order reduction. In *Proc. of the conference on Design, automation and test in Europe*, 2010.

[17] B. Bond and L. Daniel. A piecewise-linear moment-matching approach to parameterized model-order reduction for highly nonlinear systems. *IEEE Trans. Computer-Aided Design*, 26(12):2116–2129, Dec 2007.

[18] C.P. Coelho, J.R. Phillips, and L.M. Silveira. Robust rational function approximation algorithm for model generation. In *Proceedings of the 36th Design Automation Conference*, pages 207–212, New Orleans, June 1999.

[19] B. Gustavsen and A. Semlyen. Rational approximation of frequency domain responses by vector fitting. *Power Delivery, IEEE Transactions on*, 14(3):1052–1061, jul 1999.

[20] K.C. Sou, A. Megretski, and L. Daniel. A quasi-convex optimization approach to parameterized model order reduction. *IEEE Trans. Computer-Aided Design*, 27(3):456–469, march 2008.

[21] Z. Mahmood, B. Bond, T. Moselhy, A. Megretski, and L. Daniel. Passive reduced order modeling of multiport interconnects via semidefinite programming. In *Proc. of the conference on Design, automation and test in Europe*, 2010.

[22] A. Semlyen and B. Gustavsen. Vector fitting by pole relocation for the state equation approximation of nonrational transfer matrices. *Circuits, Systems, and Signal Processing*, 19(6):1531–5878, Nov 2000.

[23] B. Gustavsen. Passivity enforcement of rational models via modal perturbation. *Power Delivery, IEEE Transactions on*, 23(2):768–775, april 2008.

[24] S. Grivet-Talocia. Passivity enforcement via perturbation of hamiltonian matrices. *Circuits and Systems I: Regular Papers, IEEE Transactions on*, 51(9):1755–1769, sept. 2004.

[25] D. Deschrijver and T. Dhaene. Stability and passivity enforcement of parametric macromodels in time and frequency domain. *Microwave Theory and Techniques, IEEE Transactions on*, 56(11):2435–2441, nov. 2008.

[26] B. Bond, Z. Mahmood, Y. Li, R. Sredojevic, A. Megretski, V. Stojanovic, Y. Avniel, and L. Daniel. Compact modeling of nonlinear analog circuits using system identification via semidefinite programming and incremental stability certification. *IEEE Trans. Computer-Aided Design*, 2010.

[27] L. Ljung. *System Identification, Theory for the User*. Prentice Hall, 2nd edition, 1999.

[28] J. C. Willems. Dissipative dynamical systems. *Arch. Rational Mechanics and Analysis*, 45:321–393, 1972.

[29] A. Megretski. Convex optimization in robust identification of nonlinear feedback. In *IEEE Conference on Decision and Control*, December 2008.

[30] A. Megretski. SPOT: Systems polynomial optimization toolbox. *http://web.mit.edu/ameg/www/*.

[31] J.F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999.

[32] K.C. Sou. SQUID: A matlab tool for stable quasiconvex identification of linear systems. *http://www.rle.mit.edu/cpg/codes/squid*.

[33] B. Bond. STINS: A matlab tool for stable identification of nonlinear systems via semidefinite programming. *http://www.rle.mit.edu/cpg/codes/stins*.

[34] Z. Mahmood. PMSQUID a matlab tool for parameterized multiport quasiconvex identification of linear systems. *http://www.rle.mit.edu/cpg/codes/pmsquid*.